USAID
FROM THE AMERICAN PEOPLE

STAT REVIEW/EXECUTIVE DASHBOARD TASK ORDER

EXECUTIVE DASHBOARD SYNTHESIS OF BEST PRACTICES

October 26, 2011

This publication was produced for review by the United States Agency for International Development. It was prepared by Zaloni, Inc and Development & Training Services, Inc. (dTS).

Prepared for the United States Agency for International Development, USAID Contract Order Number AID-OAA-I-10-00014 Order Number: AID-OAA-TO-11-00045, STAT Review Executive Dashboard.

Implemented by:
Development & Training Services, Inc. (dTS)
4600 North Fairfax Drive, Suite 304
Arlington, VA 22203

Phone: +1 703-465-9388
Fax: +1 703-465-9344
www.onlinedts.com

EXECUTIVE DASHBOARD SYNTHESIS OF BEST PRACTICES

October 26, 2011

DISCLAIMER

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

CONTENTS

EXECUTIVE SUMMARY

In this report we present a synthesis of industry best practices relating to executive dashboard models. It covers the implementation approach, visualization tools, data input and data quality requirements for developing and deploying dashboards. It presents data input automation approaches, data quality assurance processes and deployment options that should be considered for building Business Performance Dashboard solutions. Data input automation is one of the key considerations to ensure correctness and currency of the data. However, given the scope of the proposed effort, it is recognized that such implementation of best practices will be beyond the scope of this project.

Based on the best practices around data integration from disparate sources, data visualization and deployment options, a set of recommendations are made that should be considered for the current and future iterations of the Executive Dashboard.

OVERVIEW OF EXECUTIVE DASHBOARD BEST PRACTICES

USAID seeks to strengthen its central management performance capability. Stronger management performance capability will result in greater accountability, consistent monitoring, and reliable reporting that will increase the Agency's ability to move toward a more efficient and effective management platform. USAID currently lacks central management performance capability resulting in weak accountability, ad-hoc monitoring, and unreliable reporting that puts the Agency at risk.

M/MPBP is launching an effort to build USAID's management performance capacity to include strategic planning, monitoring and evaluation, and communicating results. Modeled after the decades-long performance management experience on the program side, this effort will initiate long-term planning for management functions; establish rigorous standards for management metrics and data quality; and promote a process for transparent monitoring and reporting via executive dashboards.

In this document, we explore the industry best practices relating to executive dashboard models. These best practices cover the implementation approach, visualization, data input and data quality requirements for developing and deploying dashboards. It further defines stakeholder engagement for providing data for the executive dashboard, data input automation approaches, data quality assurance processes and deployment. At the outset we explore the Market trends to set the context of what others are doing in this space.

We reviewed a number of the executive dashboards implemented by local, state and federal agencies that are available on the web. Information from these dashboards is provided in the Executive Dashboard Background Research Compendium. However, the information about back-end implementation of these dashboards such as data collection mechanisms, data quality assurance, data processing, and infrastructure used to implement these dashboards is not available when reading the dashboards over the internet. Therefore, in this document we rely on the reference material that summarize the industry standards and best practices for dashboard implementation rather than direct observation of the currently implemented dashboards.

MARKET OVERVIEW

As Business Intelligence (BI) technology evolves, even basic reporting tools have not yet become commodities, and vendors are constantly surpassing each other with advanced reporting features such as high-volume scalability, Web 2.0 rich Internet application (RIA) interfaces, and many others. Today, majority of all companies are using Business Intelligence reporting solutions, and about half of these are still expanding and upgrading their reporting applications. Analytic solutions that slice and dice data from multiple sources and deliver the results as dashboards are currently implemented by almost 40% of all organizations surveyed by Forrester in a study. According to them, the next big focus area in Business Intelligence applications, will be in the category of Business Performance Solutions (BPS).

As many organizations have already implemented analytic BI solutions, they're starting to look into a next layer of BI technology that relates data from various sources to key performance indicators (KPIs), creates internal and external benchmarks, and provides decision support to improve business

performance. When talking about BPS, most people think of financial performance management. Financial BPS has the highest adoption rate of all performance management solutions, but the levels of interest in other business lines, including operational performance management in sales, production, procurement, and sustainability, are all at a similar level. As seen by Forrester, there is tremendous opportunity in the coming years for BI applications and users to turn the interest in BPS into deployments.

DATA COLLECTION

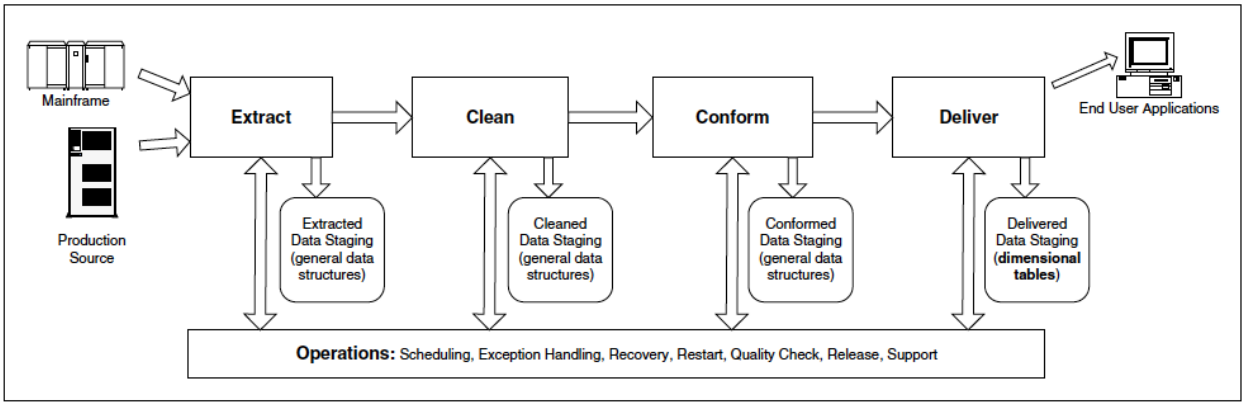
Getting the data for visualization by the Business Performance Solutions is an important step and has consequences on the implementation of all downstream processes and systems to be able to derive insights from that data. In this section, we explore the various considerations for gathering data for monitoring Key Performance Indicators in a Business Performance dashboard.

DATA INPUT AUTOMATION

Data input automation is used to extract data from the source systems, enforces data quality and consistency standards, conforms data so that separate sources can be used together, and finally delivers data in a presentation-ready format so that application developers can build applications and end users can make decisions.

Data Extraction is done via variety of ways – using standard connectors to the databases or using proprietary code or by creating flat files. After extraction, the data is transformed, or modified, depending on the specific business logic involved so that it can be sent to the target repository. There are a various ways to perform the transformation, and the work involved varies. The data may require reformatting only, but most operations also involve cleansing the data to remove duplicates and enforce consistency. The software does examine individual data fields and apply rules to consistently to convert the contents to the form required by the target repository or application. In addition, this process could involve standardizing name and address fields, verifying telephone numbers or expanding records with additional fields containing demographic information or data from other systems.

The building blocks of a system that gathers data from multiple sources and makes it available for Presentation layer can be described as follows:



ETL Workflow, Source: The Data Warehouse ETL Toolkit[1]

- **Extract**: The raw data coming from the source systems is usually written directly to disk with some minimal restructuring but before significant content transformation takes place. Data from structured source systems (such as databases, or XML data sets) often is written to flat files or relational tables in this step. This allows the original extract to be as simple and as fast as possible and allows greater flexibility to restart the extract if there is an interruption. Initially captured data can then be read multiple times as necessary to support the succeeding steps. In some cases, initially captured data is discarded after the cleaning step is completed, and in other cases data is kept as a long-term archival backup. The initially captured data may also be saved for at least one capture cycle so that the differences between successive extracts can be computed.

- **Clean**: In most cases, the level of data quality acceptable for the source systems is different from the quality required by the target systems. Data quality processing may involve many discrete steps, including checking for valid values, ensuring consistency across values, removing duplicates, and checking whether complex business rules and procedures have been enforced. Data-cleaning transformations may even involve human intervention and the exercise of judgment. The results of the data-cleaning step are often saved temporarily as an intermediate step in the Data Automation process. Data exceptions found during this step should be reported to build a case for improvements in the source system.

- **Conform**: Data conformation is required whenever two or more data sources are merged in the data warehouse. Data conformation is a significant step that is more than simple data cleaning. Data conformation requires an enterprise-wide agreement to use standardized domains and measures.

- **Deliver**: The whole point of the whole Data Input automation process is to make the data ready for querying for Business Performance Dashboards. The final step in this process is to structure the data into a set of simple, symmetric schemas known as dimensional models, or equivalently schemas so that they can be loaded into the target systems for querying. These schemas are designed in a manner so that they significantly reduce query times and simplify application development.

DATA QUALITY

Data quality is a perception or an assessment of data's fitness to serve its purpose in a given context. Aspects of data quality include:

- Accuracy
- Completeness
- Update status
- Relevance
- Consistency across data sources
- Reliability
- Appropriate presentation
- Accessibility

Within an organization, acceptable data quality is crucial to operational and transactional processes and to the reliability of Business Intelligence reporting. Data quality is affected by the way data is

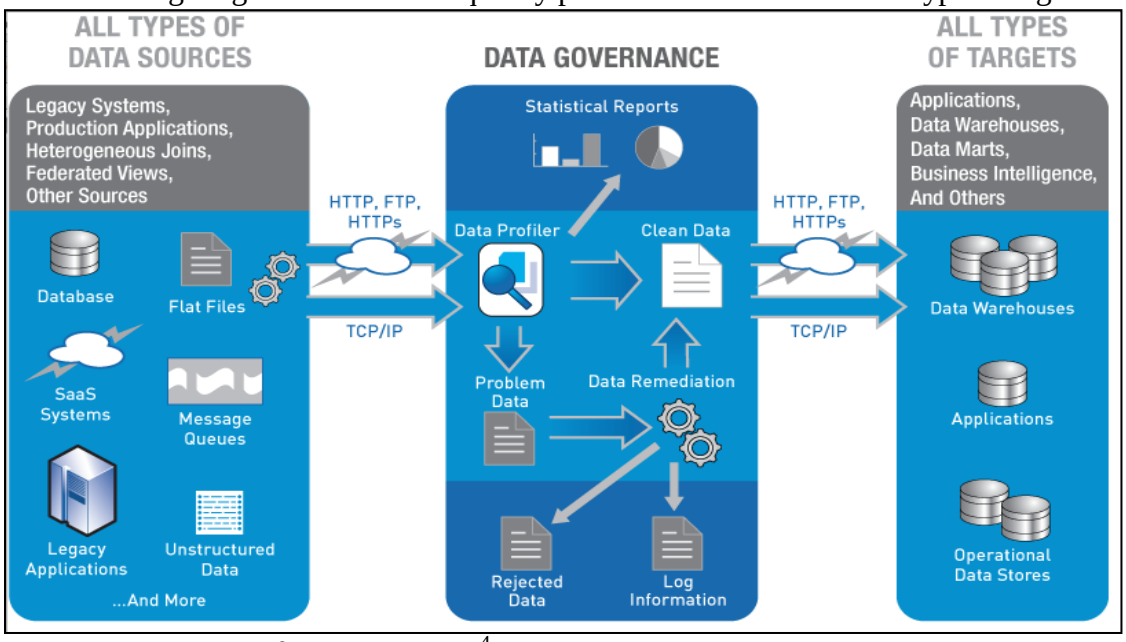
entered, stored and managed. Data quality assurance (DQA) is the process of verifying the reliability and effectiveness of data.
Maintaining data quality requires going through the data periodically and scrubbing it. Typically this involves updating it, standardizing it, and de-duplicating records to create a single view of the data, even if it is stored in multiple disparate systems.
Ralph Kimball[2], one of the leading authorities in Data Management proposes a nine-step template for any organization addressing data quality:

1. Declare a high-level commitment to a data quality culture.
2. Drive process reengineering at the executive level.
3. Invest to improve the data entry environment.
4. Invest to improve application integration.
5. Invest to change how processes work.
6. Promote end-to-end team awareness.
7. Promote inter-departmental cooperation.
8. Publicly celebrate data quality excellence.
9. Continuously measure and improve data quality.

Once the executive support and the organizational framework are ready, then specific technical solutions are appropriate. Goals for the technology include:

- Early diagnosis and triage of data quality issues
- Specific demands on source systems and integration efforts to supply better data,
- Specific descriptions of data errors expected to be encountered in extract, transform and load (ETL),
- Framework for capturing all data quality errors,
- Framework for precisely measuring data quality metrics over time, and
- Quality confidence metrics attached to final data.

The following diagram shows Data quality processes that are used in a typical organization.



Source: Pervasive Software web-site [4]

Given the best practices in Data Integration and Data Quality as pre-requisites for successful deployment of Business Performance Solutions, it is recognized that based on the current scope of the Executive Dashboard for USAID Management Bureau, the proposed implementation will be require manual user entry of the data. A data model will be defined to ensure consistent modeling of Performance indicators. Based on the data model for the Indicators, a set of forms will be provided to the users for data entry of the Indicator values. It is assumed that the user will be responsible for the data quality and correctness of the values entered so that they can be accurately displayed in the Dashboard. Future implementations should consider direct pulls of the data from USAID Enterprise Information Systems to avoid manual entry and ensure currency of the data.

PRESENTATION AND DATA VISUALIZATION

Data presentation can be simple, elegant and descriptive. There is a variety of conventional ways to visualize data – tables, histograms, pie charts and bar graphs that can be used for Business Performance Indicators. Various navigation approaches should be considered so that users can quickly get to the Indicator that they are interested in, without having to go through them sequentially. At the same time, dashboard design should consider layouts that will facilitate structured review meetings and presentations. A typical approach is the "PowerPoint" like navigation using the slide sorters to guide the user during a review.

Our approach for implementation of the presentation layer is using a Rich Internet Application for presenting the Business Performance Solutions in a Dashboard format.

A Rich Internet Application (RIA) is a Web application that has many of the characteristics of desktop application software, typically delivered either by way of a web browser , via a browser plug-in, extensive use of JavaScript, or virtual machines. Adobe Flash, JavaFX, and Microsoft Silverlight are currently the three most common platforms, based on their penetration rates of around 96%, 76%, and 66% respectively. Users generally need to install a software framework using the computer's operating system before launching the application, which typically downloads, updates, verifies and executes the RIA. This is the main differentiator from JavaScript-based alternatives like Ajax that use built-in browser functionality to implement comparable interfaces. Rich Internet Application frameworks that include even server-side frameworks, let Application developers create applications with compelling user experiences. Web standards such as HTML5 have been developed and the compliance of Web browsers with those standards is gradually improving.

Based on an intuitive, RIA implementation, Business Performance Dashboards should consider the aspects of User collaboration and Productivity when considering an implementation. This often includes features like being able to provide Annotations of the data, ability to comment on an Indicator, create actionable items in the context of an Indicator to foster collaboration amongst stakeholders and increase productivity.

DEPLOYMENT

For deploying the Business Performance Solutions, both Physical and Virtual Infrastructure deployment models can be considered. Depending on the requirements of the Solution, deployment can be mapped to a dedicated or a shared services model.

PHYSICAL INFRASTRUCTURE

Using the Internal Hardware infrastructure – Servers, Storage and Networks, Business Performance Solutions can be deployed in a Private data center using the Physical IT infrastructure. This requires dedicated Servers and Storage to host the application including the Webservers, Database servers and any other related software. Access to these applications is granted based on the Security model defined. Since the Physical IT infrastructure is controlled within the Organization's data center, Data access is more tightly controlled based on existing policies and procedures in place.

The Physical Infrastructure model is not a shared infrastructure model. It is also a fixed model – infrastructure has to be pre-provisioned and cannot be scaled up or down on demand based on traffic or SLA requirements. A change control process is typically followed to make any changes to the infrastructure. The Physical Infrastructure model requires adequate investments with a relatively longer Return on Investment.

CLOUD BASED INFRASTRUCTURE

Cloud based model offers a new supplement, consumption, and delivery model for IT services based on Internet protocols, and it typically involves provisioning of dynamically scalable and often virtualized resources. Cloud based deployment is ideal for web-based tools or applications that users can access and use through a web browser as if the programs were installed locally on their own computers.

Cloud based infrastructures deliver applications via the Internet, which are accessed from a web browser, while the business software and data are stored on servers at a remote location. Using the Cloud based model, applications can be deployed in such a way where users are charged based on the usage and they have Service Level Agreements for Application Availability. As the application needs grow, the Virtual infrastructure can be scaled up to meet changing user requirements without having to restructure the entire application.

Security is an important consideration for Cloud based deployment of the applications. Proper design and implementation of the security model is required for cloud based deployment of Business Performance Solutions. The cloud option can be implemented using Internal Private Cloud or Public Cloud such as Amazon, Rackspace and others.

RECOMMENDATIONS

Based on some of the Best practices for Dashboard design and Data capture for creating such dashboards, here are some of the recommendations for the USAID Executive Dashboard:

1. Effective dashboards display the important data that someone must monitor to do a job in a way that can be read and understood quickly and easily. This should be done in a way where the Dashboard design avoids clutter.

2. Whenever data is displayed visually, as in a graph, minimize visual content that doesn't represent actual data. Unless non-data content supports the data in some essential way, it is nothing more than material that viewers must scan without reward. Especially on a dashboard where a great deal of data must be communicated in an instant, there is no room for meaningless visual content.

3. Avoid variety in the choice of display mechanisms. When deciding how to best display each section of data, consider the display that will communicate this data most effectively and efficiently.

4. Given the purpose of a dashboard is to monitor some aspect of the business operations, current data is usually of greater interest than historical data. It is often quite useful to display some history as meaningful context for the current measures, but the current measures ought to stand out more prominently than the history.

5. Data quality should be considered upfront. Define a model for ensuring quality assurance of the data that is entered into the Solution for displaying various Business Performance Indicators. Some of the measures can be implemented in the Source of Truth systems so that data that is being fed to the downstream applications is of the desired quality.

6. Establish a process to monitor data quality issues and track them to resolution.

7. Automate the Data entry process as much as possible. This could include API level integration with the source systems and ant ETL process to consume the data. Data could be imported in fixed intervals (e.g.: weekly, monthly).

8. Sensitive Business data needs to be securely maintained and displayed to the users based on their privilege levels. Use of a Role Based Access Control mechanism should be considered.

9. For deployment of the dashboard, consider a Shared Services model using Cloud based Infrastructure for more efficient deployment of the solution. Using the cloud-based model also enables the ability to Scale up or down based on the user demand.

10. Establish a backup and disaster recovery plan upfront and implement it for the Dashboard deployment.

REFERENCES

1. Ralph Kimball and Joe Caserta, "The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data", Indianapolis, IN: Wiley, August 2004, ISBN: 978-0-7645-6757-5
2. Ralph Kimball, "An Architecture for Data Quality", Information Management Magazine, October 2007
3. James Ward , "What is a Rich Internet Application?", James Ward personal website, October 2007, http://www.jamesward.com/2007/10/17/what-is-a-rich-internet-application/
4. Know your data quality, Pervasive Software web-site, http://www.pervasiveintegration.com/products/Pages/data_quality_tools.aspx